

# Summary of Findings

## 1. Dataset Overview

- The Titanic dataset contains **891 passengers** with a mixture of categorical (Sex, Pclass, Embarked) and numerical (Age, Fare, SibSp, Parch) features.
  - Missing values were found mainly in Age, Cabin, and Embarked columns.
- 

## 2. Univariate Analysis

- **Sex:** Approximately **65% of passengers were male**.
  - **Pclass:** Most passengers traveled in **3rd class**.
  - **Age:** Majority were **young adults** aged between 20–30 years.
  - **Fare:** Fare distribution was **highly skewed**, with most fares in the lower range.
- 

## 3. Bivariate Analysis

- **Sex and Survival:** **Females had a much higher survival rate** compared to males.
  - **Pclass and Survival:** **1st class passengers** had the highest survival rates.
  - **Embarked and Survival:** Passengers embarked from **Cherbourg (C)** had slightly better survival rates.
  - **Age and Survival:** **Younger passengers** had better survival chances.
  - **Fare and Survival:** Passengers who **paid higher fares** were more likely to survive.
- 

## 4. Correlation Analysis

- **Fare** is **positively correlated** with **Survival**.
  - **Pclass** is **negatively correlated** with **Survival**.
  - Strong negative correlation between **Fare and Pclass** (higher class passengers paid more).
-

### Key Insights

- **Gender, Class, and Fare** are major factors influencing survival.
- **Younger, wealthier, and higher-class passengers** had better chances of surviving.
- **Embarkation point** (Cherbourg) also showed some impact on survival.